By Dr. Farah
Telegram Id- https://t.me/ugcnetpsychologyforall

Factor Analysis: Assumptions, Methods, Rotation and Interpretation

Factor

Essentially a factor is a dimension or construct which is a condensed statement of the relationships between a set of variables. This has been more precisely refined by Royce (1963), who states that a factor is a construct operationally defined by its factor loadings.

This has been more precisely refined by Royce (1963), who states that a factor is a construct operationally defined by its factor loadings.

Factor loadings are the correlations of a variable with a factor.

A factor loading for a variable is a measure of how much the variable contributes to the factor; thus, high factor loading scores indicate that the dimensions of the factors are better accounted for by the variables

Common variance. This is the proportion of the variance which can be explained by common factors.

In many methods of factor analysis, the goal of extraction is to remove as much common variance in the first factor as possible (Child, 2006). The communality is the variance in the observed variables which are accounted for by a common factor or common variance

Specific variance. This is the variance which is particular to a test or variable. In the case of a test, for example, specific variance

SHOUT
SHARE
IT

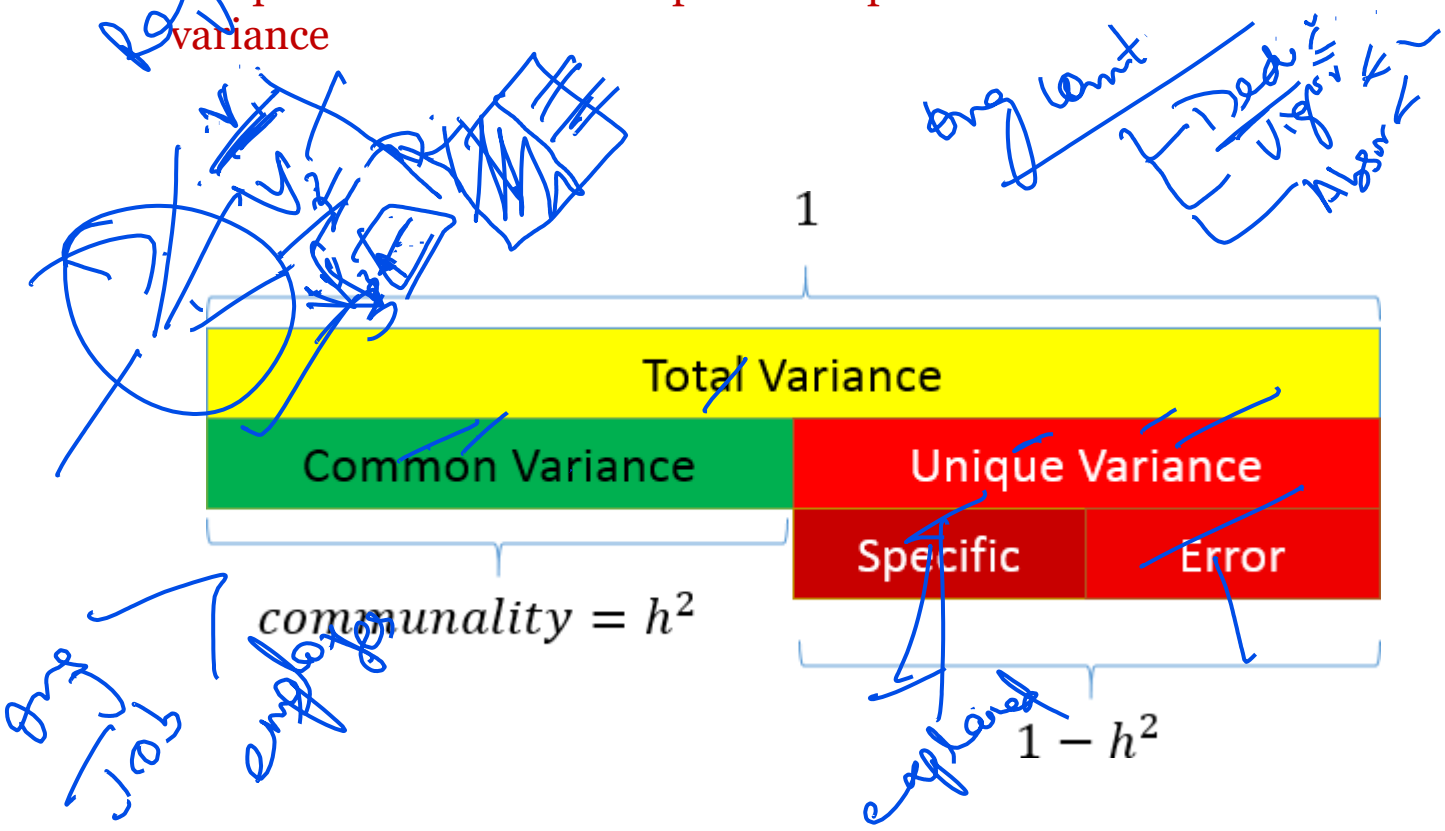
can arise from the particular form of the items in the test, especially if they are different from those in other tests, and from the particular content.

Error variance. Almost all measurement, not only in psychology, is subject to random error. With psychological tests the reliability, the correlation of the test with itself or its internal consistency, is an index of measurement error. Squaring this correlation and subtracting from 100 gives the amount of error variance. Thus, if a test has a reliability of O. 9, there is 19 per cent error variance in the scores.

The unique variance is denoted by u 2 and is the proportion of the variance that excludes the common factor variance which is represented by the formula (Child, 2006).

U (unique variance) = h2 - 1

Unique variance can be split into specific variance and error variance



Assumptions-

- The recommended sample size is at least 300 participants, and the variables that are subjected to factor analysis each should have at least 5 to 10 observations (Comrey & Lee, 1992).
- It should have at-least three variables
- It is also recommended that a heterogeneous sample is used rather than a homogeneous sample as homogeneous samples lower the variance and factor loadings
- Next, the correlation r must be .30 or greater since anything lower would suggest a really weak relationship between the variables (Tabachnick & Fidell, 2007).
- Also, a determining factor is based on the assumption that there is a linear relationship between the factors and the variables when computing the correlations (Gorsuch, 1983).
- It is important that to check for an absence of multicollinearity and singularity within the dataset by looking at the Squared Multiple Correlation

Types of Factor analysis

1. Exploratory Factor analysis
2. Confirmatory Factor analysis

Exploratory Factor Analysis

In exploratory analysis the aim is to explore the field, to discover the main constructs or dimensions. It was for this purpose that factor analysis was originally developed by Spearman (1904), in the area of human abilities. In general, in exploratory analysis the rule is to put in as many variables as possible and see what loads on the relevant factor. If our interest is in achievement at school, we simply look for all the variables which load on the same factor as the achievement measures. If our interest were more specific, say science performance, then we would look for a factor specific to that and see what variables loaded on it.

Confirmatory Factor analysis

Hypotheses testing is done by a method developed by Joreskog (1973) called confirmatory analysis. In this method, based upon previous studies or on relevant theory, factor loadings for the variables are hypothesized. Confirmatory factor analysis then proceeds to fit these loadings in the target matrix, as it is called, as closely as possible. How good the fit is can also be measured?

Factor Extraction Method

Principal Component method

The aim of principal components analysis is to be able to estimate the correlation matrix and this can be done by finding the characteristic equation of the matrix. This requires two sets of values:

A characteristic vector is a column of weights each applicable to one of the variables in the matrix. Thus, if there were six variables there would be six weights in the first vector - Va, Vb and so on. The corresponding factor loadings, Fa, Fb , can be

obtained by multiplying each element of the vector by the square root of the corresponding eigenvalue.

Definition of characteristic root or eigenvalue- The sum of squares of the factor loadings of each factor reflects the proportion of variance explained by each factor. This total amount of variance is the characteristic root or eigenvalue I, for the factor. The larger the eigenvalue the more variance is explained by the factor.

Components are real factors because they are derived directly from the correlation matrix. Common factors of factor analysis are hypothetical because they are estimated from the data. the principal axes method of factor analysis is identical to that of principal components except that instead of unity in the diagonals some other estimate of communality is inserted. This means that while the principal components method explains all the variance in a matrix the principal axes method does not.

Common Factor Analysis

Cattell (1978) prefers to work with common factors rather than components even though in large matrices differences between the two solutions are trivial (Harman 1976).

There are several important differences.

Components are real factors. In the principal component method, the factors are actual combinations of variables. The factor loadings are the correlations of these combinations with the factors.

Common factors are hypothetical. Common factors have to be estimated from actual variables and to obtain them mathematical procedures must be used which specify factors in terms of common variance.

Communalities must be less than unity. Because factors are concerned with common variance, it follows that the communalities, the diagonal elements of the correlation matrix, must be less than one.

Common and unique variances are not separated in PCA. When ls are placed in the diagonals, as is the case with principal components analysis, this separation is not complete. This is because the variables themselves determine the factors but the variables are composed of unique and common variance. In common factor analysis unique and common variance are separated.

Both methods try to reduce the dimensionality of the dataset down to fewer unobserved variables

The other main difference between PCA and factor analysis lies in the goal of your analysis. If your goal is to simply reduce your variable list down into a linear combination of smaller components then PCA is the way to go. However, if you believe there is some latent construct that defines the interrelationship among items, then factor analysis may be more appropriate

Maximum Likelihood method

The main concept to know is that ML also assumes a common factor analysis using the R2 to obtain initial estimates of the communalities, but uses a different iterative process to obtain the extraction solution.

Although the initial communalities are the same between PAF and ML, the final extraction loadings will be different, which means you will have different Communalities, Total Variance Explained, and Factor Matrix tables (although Initial columns will overlap). The other main difference is that you will obtain a Goodness-of-fit Test table, which gives you an absolute test of model fit. Non-significant values suggest a good fitting model.

Principal Axis Factor method

A method of extracting factors from the original correlation matrix, with squared multiple correlation coefficients placed in the diagonal as initial estimates of the communalities. These factor loadings are used to estimate new communalities that replace the old communality estimates in the diagonal. Iterations continue until the changes in the communalities from one iteration to the next satisfy the convergence criterion for extraction.

PCA	EFA	CFA
Reduces items to subsets	Reduces items to subsets	Reduces items to subsets
Descriptive technique	Modelling technique	Modelling technique
Exploratory Approach	Exploratory Approach	Confirmatory Approach
Analysis of variances	Analysis of covariances	Analysis of covariances
Less appropriate for measurement error	More appropriate for measurement error (Latent constructs)	More appropriate for measurement error (Latent constructs)
Linear combinations of items- total variance	Extracted Factors- explain relationship, common variance	Extracted Factors- explain relationship, common variance

Rotation

Orthogonal (uncorrelated, 90°)

Varimax- Simplify the number of variables

Quartimax- Simplify the numbers of factors

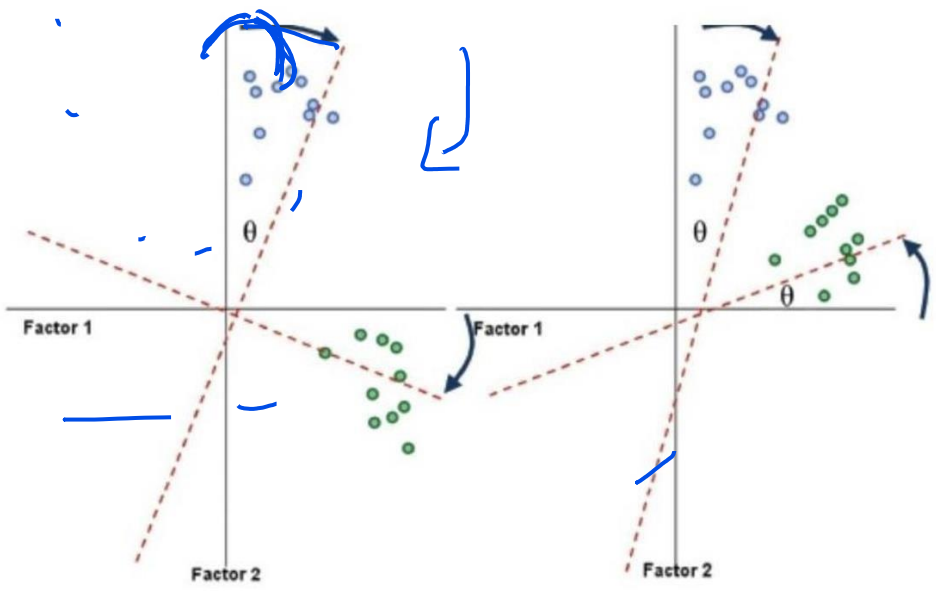
Equamax-

Oblique – Correlated

Oblimin- simplify the structure and mathematics of the output

Quartimin

Promax- expedient because of its speed in larger data set.



Orthogonal Rotation Oblique Rotation

Number of factors to be retained

The eigenvalues and scree test (i.e., scree plot) are used to determine how many factors to retain. One criterion that can be used to determine the number of factors to retain is Kaiser's criterion which is a rule of thumb. This criterion suggests retaining all factors that are above the eigenvalue of 1 (Kaiser, 1960). Another criterion is based on Jolliffe's criterion which recommends retaining factors above .70 (Jolliffe, 1986).